# 内容聚合系统的几点构思

目前河南有线正在规划内容聚合系统，收集存储EPG信息、已上线资产元数据、用户行为数据，希望对这些数据进行相关性分析，向用户提供精准产品（影片）推荐、直播节目推荐和搜索功能。

## 系统架构



内容聚合系统从外部能力系统获取数据，如下。

视频类：

1. 从直播频道内容自动识别系统获取节目拆条信息；
2. 从双向电视业务框架子系统获取内容元数据，这些内容元数据描述了存储在Cable和OTT推流系统的影片资产信息；
3. 从第三方的独立OTT业务获取内容元数据，内容元数据的扩展信息和其他系统结合后，推送给用户。（如CNTV）
4. 从网络爬虫系统获取网页和视频数据；
5. 从收拾分析系统中获取资产点播和节目收视数据；
6. 从网盘系统中获取视频元数据；
7. 从EPG系统中获取直播频道的节目单；
8. 从民生查询及缴费系统中获取业务接口。
9. 把汇聚的网页图文信息，直播和ott视频元数据整合起来，快速形成具有模板、策划的专题进行业务推送和推荐。
10. 汇聚的内容和用户账户关联。
11. 内容整合后的搜索和推荐。

内容聚合系统对上述数据进行相关分析后，向用户提供智能EPG服务、搜索和推荐服务。

在能力系统上，目前河南已经部署了双向电视业务框架子系统、收视分析系统、EPG系统，门户系统和推流系统。尚缺消息引擎系统、直播频道内容自动识别系统、网络爬取系统、网盘系统、民生查询及缴费系统。

## 数据库要求

采用分布式数据库和非结构化数据库，降低存储成本和满足未来不确定数据结构的需求。